US005352588A

United States Patent [19]

Fischetti et al.

[11] **Patent Number: 5,352,588**

[45] **Date of Patent: Oct. 4, 1994**

[54] **STREPTOCOCCAL IMMUNOGLOBULIN A BINDING PROTEIN ENCODED BY EMML2.2**

[75] Inventors: **Vincent A. Fischetti**, West Hempstead; **Debra E. Bessen**, New York, both of N.Y.

[73] Assignee: **Rockefeller University**, New York, N.Y.

[21] Appl. No.: **813,584**

[22] Filed: **Dec. 24, 1991**

[51] **Int. Cl.**[5] **C07K 13/00**; C12N 15/31; C12N 1/21

[52] **U.S. Cl.** **435/69.1**; 530/350; 536/23.7; 435/320.1; 435/252.3; 435/252.33

[58] **Field of Search** 536/23.7; 530/350; 435/252.3, 252.33, 320.1, 69.1

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,757,134	7/1988	Blake et al.	530/350
5,210,183	5/1993	Lindahl et al.	530/350

OTHER PUBLICATIONS

Bessen, D. E. et al. (1992) *Infect. Immun.* 60: 124–35.
Jeppson, H., et al. (1992) *FEMS Microbiol. Lett.* 92: 139–46.
Frithz, E., et al. (1989) *Mol. Microbiol.* 3: 1111–19.
Haanes, E. J., et al. (1989) *J. Bacteriol.* 171: 6397–6408.
Khandke, K. M., et al. (1988) *J. Biol. Chem.* 263: 5075–82.

Primary Examiner—Garnette D. Draper
Assistant Examiner—David L. Fitzgerald
Attorney, Agent, or Firm—Saliwanchik & Saliwanchik

[57] **ABSTRACT**

The subject invention concerns a novel polynucleotide sequence cloned from emm2.2 gene of a Group A streptococcus, Type II strain which codes for an IgA-binding protein, ML2.2. A process for producing the protein is given. The invention also concerns the protein in an immunoadsorbent and as a tracer for use in measuring and purifying IgA. Kits are given comprising the immunoadsorbent and the tracer form of the protein.

7 Claims, 2 Drawing Sheets

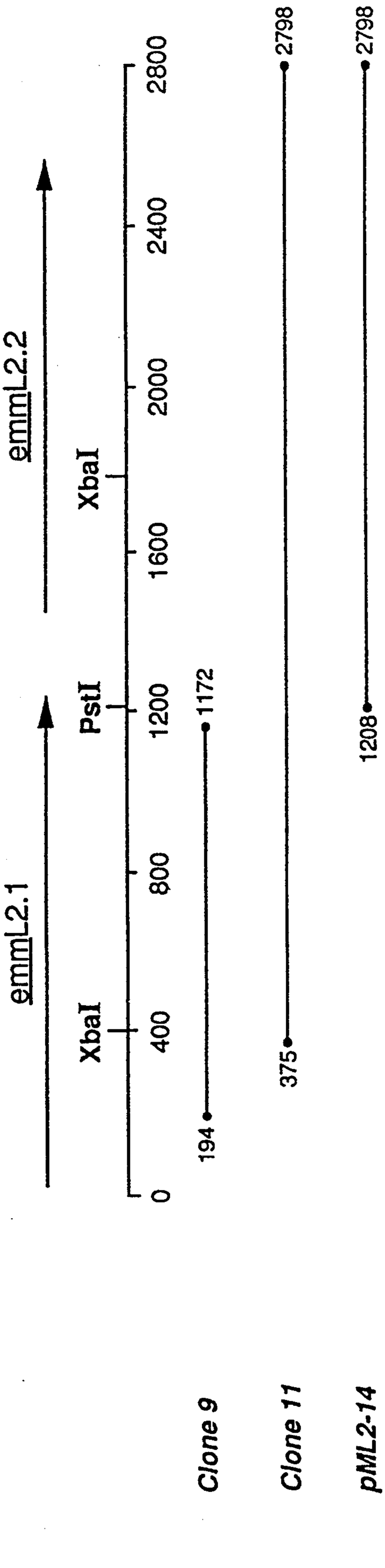
emmL2.2
XbaI
1600
2000
2400
2800
2798
2798
emmL2.1
XbaI
PstI
0
400
800
1200
194
1172
375
1208
209
Clone 9
Clone 11
pML2-14
emmL2.1 mRNA
1

Figure 1

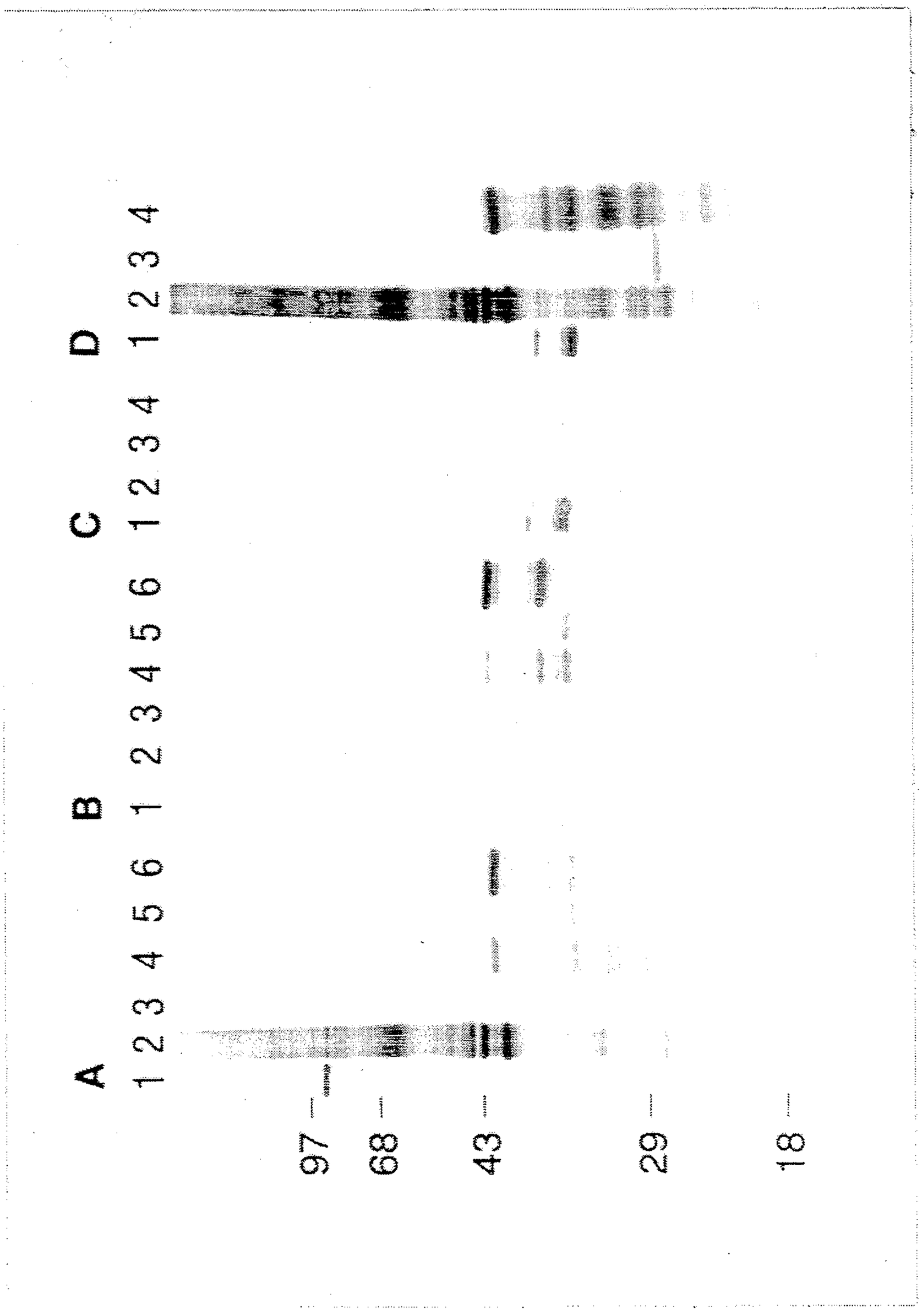

Figure 2

STREPTOCOCCAL IMMUNOGLOBULIN A BINDING PROTEIN ENCODED BY EMML2.2

This invention concerns a gene coding for a novel immunoglobulin A (IgA) binding protein from group A streptococci of Class II, a process for producing the protein by genetic recombination and the use of the protein to capture and measure IgA.

Nucleotide accession number: the nucleotide sequence encompassing the gene emmL2.2 coding for the IgA binding protein of the present invention is available from EMBL/GenBank/DDJB under accession number X61276. DNA sequence for the polynucleotide coding for protein ML2.2 is submitted on a floppy disk with the present specification.

BACKGROUND OF THE INVENTION

Group A streptococci are responsible for a wide variety of human diseases, the most common of which are nasopharyngitis and impetigo. Nearly all clinical isolates have the antiphagocytic factor, M protein, on their surface. This virulence factor displays extreme antigenic diversity within its amino-terminal region. It is these highly variable portions of M proteins which form the basis of the serological typing scheme which was formulated in the 1930s prior to knowledge of any structural detail (Lancefield, R. C., *J. Immunol.* 89:307–313 (1962)).

In recent years, the sequences of M or M-like proteins have been reported (Frithz, E., Heden, L. O. and Lindahl, G. *Mol. Microbiol.* 3:1111 (1989)). The M-like molecules are structurally similar to M proteins in that they exhibit significant levels of sequence homology; however, they are not considered to be M protein itself because an antiphagocytic property has not been formally demonstrated.

Group A streptococci can be divided into two major classes partly on the basis of their immunoreactivity with a pair of monoclonal antibodies directed to epitopes which lie within the relatively conserved half of M proteins (Bessen, D. and Fischetti, V. A. *J. Exp. Med.* 172:1757 (1990); Bessen, D., Jones, K. F. and Fischetti, V. A. *J. Exp. Med.* 169:269 (1989)). Class I isolates are defined as those binding one or both monoclonal antibodies, whereas class II isolates do not bind either monoclonal antibody. In addition, the classes differ in their ability to exhibit opacity factor activity, and in several pathogenic properties of these organisms (Bessen, D. and Fischetti, V. A. *J. Infect. Dis.* 161:747 (1990)). For example, nearly all serotypes found in association with major outbreaks of rheumatic fever are class I. The classes also differ in ability to bind IgA, with this activity being specific for class II. Obviously this property is important to the function of streptococci as the arp4 protein reported by Lindahl et al. (European Patent Application 367890), e.g., is from class II streptococci and the IgA-binding protein of Russell-Jones et al. (U.S. Pat. No. 4,757,134) is from group B streptococci.

Immunoglobulin A is an important component of the bodily response to pathogens and other disorders. An early mucosal immune response to invasion is production of antibodies of the A class and as a result, elevated levels of IgA can be found in extracts of infected membranes such as saliva, urine, feces and urogenital extracts. Certain disorders such as kidney or liver malfunctions and early cancer detection may also be correlated to IgA. Indeed the human body produces more IgA daily than any other antibody class.

Accordingly, reagents to capture and measure IgA have been sought.

SUMMARY OF THE INVENTION

Two M or M-like protein genes were cloned from a single streptococcus A class II isolate. It was discovered that the product of the downstream gene, hereinafter termed ML2.2, is a protein that exhibits IgA-binding activity (Bessen, D. E. and Fischetti, V. A., *Infection and Immunity* (1992) "Nucleotide Sequence of Two Adjacent M or M-Like Protein Genes of Group A Streptococci: Different RNA Transcript Levels and Identification of a Unique IgA Binding Protein," 60:124–135, herein incorporated by reference). Described here is a novel process for producing high quantities of this IgA binding protein hereinafter referred to as ML2.2.

The polynucleotide of the present invention comprises DNA of approximately 1.6 kb which codes for a polypeptide having the ability to bind IgA. The polypeptide is an approximately 36 to 42 kD protein expressed by a Group A streptococcus of the IIb class, or a fragment or equivalent of the protein. The polynucleotide has the DNA sequence given in FIG. **1**. The amino acid sequence of the protein product of the polynucleotide is also given in FIG. **1**. Most especially the present invention sets forth the polynucleotide sequence which codes for a polypeptide which binds IgA.

The present invention further concerns a plasmid comprising this polynucleotide sequence, preferably plasmid pML2-14.

According to the process of the subject invention, microorganisms which have been transformed with the gene coding for ML2.2 produce and secrete large quantities of the recombinant protein. Specifically, according to the subject invention, a suitable host, an *Escherichia coli,* for example, can be transformed with 1.6 kb DNA comprising the nucleotide sequence shown in SEQ ID NO. 1. This sequence codes for the IgA binding protein of approximately 42,000 daltons designated ML2.2, whose amino acid sequence is also shown in SEQ ID NOS. 1 and 3.

The product of this process, ML2.2, may be used to detect the presence of IgA in a biological sample. To detect IgA, the sample is contacted with the protein under conditions for suitable for reaction, in the presence of suitable buffers and salts, for example, to bind IgA to the protein and the bound product is observed.

Most preferably the protein of the present invention is affixed to a solid surface, a microtiter plate, for example, or an inert bead such as polystyrene or latex to form an immunoadsorbent for removing IgA from a biological sample. IgA may be removed from a sample by contacting the sample with the immunoadsorbent under conditions which allow IgA to bind to the immunoadsorbent. Microtiter plates, for example, may be coated with the protein or beads may be packed into a column through which sample may be passed.

The present invention also sets forth a process for preparing essentially pure IgA from a biological sample, serum, for example, containing IgA. In this process the sample is then contacted with the immunoadsorbent comprising the protein of the present invention, allowing IgA in the sample to bind to the immunoadsorbent. The bound IgA is washed, preferably with a suitable buffer, to remove associated proteins and other contam-

inants. Finally IgA itself is eluted from the immunoadsorbent by buffers, electroelution or other method known to the art.

In preferred embodiments of the present invention, the protein is conjugated to a group suitable for being observed. Any such groups known to persons skilled in the an may be used. Enzymes such as alkaline phosphatase or peroxidase, metals such as gold or a member of the biotin-avidin binding pair may be used as labels.

The present invention also sets forth kits comprising immunoadsorbent wherein the protein of the present invention is affixed. Kits may also comprise the protein of the present invention labeled for use as tracers.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** gives the region of the T2/MR chromosome cloned and sequenced to produce the nucleotide sequence of SEQ ID NO. 1. Arrows indicate the coding regions of emmL2.1 and emmL2.2. Positions of streptococcal DNA that were cloned into the M13mp19 (inserts from clones 9 and 11) and pUC18 (pML2-14) vectors used for sequencing are shown.

FIG. **2** gives the Western blot analysis of the cloned emmL2.2 gene product. Lanes 1 contain lysin extract of 29452, a streptococcal strain (type 22, class II) obtained from the Institute of Hygiene and Epidemiology, Prague, Czechoslovakia, Lanes 2 contain lysin extract of T2/MR, a M-rich isolate of Group A streptococcal strain T2/44/RB4/119 (Class II) from the Lancefield collection (The Rockefeller University), Lanes 3 contain whole *E. coli* XL-1 cells and Lanes 4 contain whole *E. coli* XL-1 cells harboring pML2-14. Blot A is anti-coli6, blot B is human myeloma IgA, blot C is human IgG-Fc fragment and blot D is anti-peptide d240 to 260. The positions of molecular size markers (in kD) are shown to the left.

DETAILS OF THE INVENTION

This invention provides a novel recombinant protein and a novel gene and methods for producing this protein. The novel recombinant protein, and subfragments thereof, have affinity for IgA and can be used in a variety of assays and kits.

One aspect of the subject invention is a gene coding for a recombinant protein. The nucleotide sequence of this gene is shown in SEQ ID NO. 1. SEQ ID NOS. 1 and 3 also show the deduced amino acid sequence of the recombinant protein encoded by the gene shown in SEQ ID NO. 1.

The invention further concerns a recombinant polynucleotide sequence comprising a vector in which a DNA sequence coding for the subject recombinant protein, or a fragment thereof, expressible in a suitable host has been inserted. Thus, said vector encodes the novel IgA binding protein and/or a fragment of this protein with substantially the same binding properties to immunoglobulin A. Specifically, the vector may be chosen from plasmids, phage DNA, or derivatives or fragments thereto, or combinations of plasmids and phage DNA and yeast plasmids.

The invention also concerns a host infected, transformed, or transfected with a recombinant DNA molecule comprising a vector in which a DNA sequence coding for the desired protein, or fragment thereof, expressible is a suitable host has been inserted. The inserted DNA is characterized in that the DNA sequence codes for the recombinant IgA binding protein and/or a fragment of this protein with substantially the same binding properties to immunoglobulin A. Among the many suitable hosts that can be infected, transformed, or transfected with the recombinant DNA molecule according to the invention and thereby express this protein or fragments thereof are gram positive or negative bacteria such as *E. coli, Bacillus subtilis,* insect cells and yeast cells.

The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. These procedures are all described in Maniatis, T., E. F. Fritsch, and J. Sambrook (1982) *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, New York. Thus, it is within the skill of those in the genetic engineering art to extract DNA from microbial cells, perform restriction enzyme digestions, electrophorese DNA fragments, tail and anneal plasmid and insert DNA, ligate DNA, transform cells, e.g., *E. coli* cells, prepare plasmid DNA, electrophorese proteins, and sequence DNA.

The DNA sequence of the subject invention can be most readily obtained by a person skilled in the art by isolating said DNA from Group A streptococcal strain T2/44/RB4/119, the M2 typing strain from the Lancefield collection (The Rockefeller University). The nucleotide sequences disclosed herein can also be prepared by a "gene machine" by procedures well known in the art. This is possible because of the disclosure of the nucleotide sequence.

As is well known in the art, the amino acid sequence of a protein is determined by the nucleotide sequence of the DNA. Because of the redundancy of the genetic code, i.e., more than one coding nucleotide triplet (codon) can be used for most of the amino acids used to make proteins, different nucleotide sequences can code for a particular amino acid. The amino acid sequences of the recombinant IgA binding protein and fragments therefore, of the subject invention can be prepared by nucleotide sequences other than that which is shown in SEQ ID NO. 1. Functionally equivalent nucleotide sequences encoding the novel amino acid sequence of these proteins and fragments can be prepared by known synthetic procedures. Accordingly, the subject invention includes such functionally equivalent nucleotide sequences.

Thus the scope of the subject invention includes not only the specific nucleotide sequences depicted herein, but also all equivalent nucleotide sequences coding for molecules with substantially the same antigenic, immunogenic, or therapeutic activity.

Further, the scope of the subject invention is intended to cover not only the specific amino acid sequences disclosed, but also similar sequences of proteins or protein fragments having comparable biological activity.

Preparation of nucleotide. Randomly sheared chromosomal DNA, derived from a low hemolytic colony of strain T2/44/RB4/119 termed T2/MR (type 2, class II) was cloned into lambda gt11 through EcoRI linkers and partially sequenced (2). Plaques were first screened with anti-ColiM6 and then with M2 typing sera. Two clones were obtained: clones 9 and 11. Purified lambda gt11 replicative-form DNA containing inserts were subcloned into both M13mp19 and pUC18 vectors. The insert from clone 11 was subcloned into pUC18 to generate pML2-11. The 1.6 kb PstI-EcoRI fragment derived from pML2-11 was ligated into pUC18 to construct pML2-14.

Purification of polypeptide. Purification of a 34.5 kD fragment of ML2.2, containing IgA-binding activity

was performed by growing *E. coli* containing pML2-14 to mid-log, and preparing a periplasmic fraction. The fraction was dialyzed in buffer containing protease inhibitors and then contacted with MonoQ and washed in 0.05 M Tris (pH 8) containing 0.005 M EDTA. The protein was in fall-through fraction.

The protein was affinity purified with immobilized human myeloma IgA and then eluted with glycine, pH 2.0. The first eluted fractions were re-applied to the column. Eluants were dialyzed and concentrated. Final purification was obtained by contacting eluant with mono-S in 0.05 M NaOAc, pH 5.5 and 0.005 M EDTA. Elution was obtained in buffer containing gradient up to 1.0 M NaCl.

DNA sequencing. Foreign DNA cloned into M13 and pUC vectors was sequenced by the dideoxy-chain termination method. Overlapping inserts were generated in M13mp19 by T4 polymerase digestion. Streptococcal DNA cloned into pUC18 vectors was sequenced because of the inability to clone the sense strand of major portions of insert 11 into M13 bacteriophage. The DNA sequence of pML2-14 is given in SEQ ID NO. 1.

The longest open reading frame for the emmL2.2 gene extends from nucleotide 1452 to 2567. The first 41 amino acids are homologous to signal peptides of other M and M-like proteins (12, 14). Therefore, the mature emmL2.2 gene product is expected to be 331 residues with a predicted molecular weight of 36,769; this size is in reasonable accordance with the 42 kD band expressed by *E. coli* harboring pML2-14 (FIG. **1**, Lanes 4). There is a single region of sequence repeats in the ML2.2 protein, consisting of three 23-residue C repeat segments separated by two spacers (spanning amino acids 90 to 203).

Sequence identities. Amino acids 1 through 71 of the mature ML2.2 protein sequence display only very limited homologies to the ML2.1 protein located upstream or with other M and M-like molecules. According to maximal alignment of sequences by the Protalign algorithm, there is 53% amino acid sequence identity between ML2.1 and ML2.2 proteins, and the homology is located for the most part within their carboxy-terminal two-thirds. ML2.2 protein exhibits 82% homology to the deduced sequence of ennX, an M-protein of unknown function (16) and appears to be transcriptionally silent in streptococci (6). Arp4 and ML2.2, both IgA-binding proteins, are only 56% identical in sequence, despite their similar functions.

Binding properties of ML2.2 protein.

FIG. **3** shows that whole cell lysates of *E. coli* harboring pML2-14 bind human myeloma IgA by Western blotting (Blot B, Lane 4). The gene product expressed by pML2-14 displayed several bands, many of which are likely degradation products of the major band at 42 kilodalton. Whole *E. coli* cells having no pML2-14 do not bind IgA. Also the lysin extract of streptococcal strain 29452 does not bind IgA although intact cells do bind IgA.

The emmL2.2 gene product, ML2.2 protein, fails to bind the class I-specific monoclonal antibodies which recognize epitopes in the C repeat regions of class I molecules.

The Ig-binding sites within M and M-like molecules have not been identified. The three M or M-like immunoglobulin binding proteins cloned and sequenced to date (Arp4, ProtH and ML2.2) have in common class II-like C repeat regions (2) and subtleties within the class II C repeat may influence Ig-binding.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 3

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1560 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: DNA (genomic)

(i x) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 2..40

(i x) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 252..1367

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
A TCT GCA GGT ATG CTT GCT CTA AAA CGC AAA GAA GAA AAC TAAGCATTAG      50

  Ser Ala Gly Met Leu Ala Leu Lys Arg Lys Glu Glu Asn
    1               5                  10

ACTGATGCTA AAGCTAAGAG AGAATCAAAT GATTCTCTCT TTTTGAGTGG CTAAGTAACT    110

AACAATCTCA GTTAGACCAA AAAATGGGAA TGGTTCAAAA AGCTGGCCTT TACTCCTTTT    170

GATTAACCAT ATATAATAAA AACATTAGGA AAATAATAGT AATATTAAGT TTGTTTCCTC    230

AATAAAATCA AGGAGTAGAT A ATG GCT AGA CAA CAA ACC AAG AAA AAT TAT       281
```

-continued

```
                                    Met Ala Arg Gln Gln Thr Lys Lys Asn Tyr
                                     1               5                  10

TCA CTA CGA AAA CTA AAA ACC GGT ACG GCT TCA GTA GCC GTT GCT TTG          329
Ser Leu Arg Lys Leu Lys Thr Gly Thr Ala Ser Val Ala Val Ala Leu
                15                  20                  25

ACC GTT TTG GGC GCA GGT TTT GCA AAC CAA ACA GAA GTA AGA GCT GAT          377
Thr Val Leu Gly Ala Gly Phe Ala Asn Gln Thr Glu Val Arg Ala Asp
            30                  35                  40

GAA GCT AAA AAA ATG GAA GTA AAA GAA AGT GAA AAA GAG TCC CAG TAT          425
Glu Ala Lys Lys Met Glu Val Lys Glu Ser Glu Lys Glu Ser Gln Tyr
        45                  50                  55

AAG ACG TTG GCT TTA AGA GGT GAA AAT GCT GAC CTT AGA AAT GTA AAT          473
Lys Thr Leu Ala Leu Arg Gly Glu Asn Ala Asp Leu Arg Asn Val Asn
    60                  65                  70

GCA AAA TAT TTA GAG AAA ATT AAC GCA GAA GAA GAA AAA AAT AAA AAG          521
Ala Lys Tyr Leu Glu Lys Ile Asn Ala Glu Glu Glu Lys Asn Lys Lys
75                  80                  85                  90

CTT GAA GCA ATT AAT AAA GAG CTA AAT GAG AAT TAT TAC AAA TTA CAG          569
Leu Glu Ala Ile Asn Lys Glu Leu Asn Glu Asn Tyr Tyr Lys Leu Gln
                95                  100                 105

GAT GGC ATT GAT GCT CTA GAA AAA GAA AAA GAA GAT CTC AAA ACA ACT          617
Asp Gly Ile Asp Ala Leu Glu Lys Glu Lys Glu Asp Leu Lys Thr Thr
            110                 115                 120

TTA GCT AAG ACT ACT AAA GAA AAT GAG ATT TCA GAA GCT AGC CGT AAA          665
Leu Ala Lys Thr Thr Lys Glu Asn Glu Ile Ser Glu Ala Ser Arg Lys
        125                 130                 135

GGG TTA AGC CGA GAC TTA GAA GCT TCT CGT ACA GCT AAA AAA GAG CTA          713
Gly Leu Ser Arg Asp Leu Glu Ala Ser Arg Thr Ala Lys Lys Glu Leu
    140                 145                 150

GAA GCT AAG CAT CAA AAA TTA GAA GCA GAA AAC AAA AAA CTA ACA GAA          761
Glu Ala Lys His Gln Lys Leu Glu Ala Glu Asn Lys Lys Leu Thr Glu
155                 160                 165                 170

GGC AAT CAG GTT TCA GAA GCT AGT CGT AAA GGT CTA AGT AAC GAC TTA          809
Gly Asn Gln Val Ser Glu Ala Ser Arg Lys Gly Leu Ser Asn Asp Leu
                175                 180                 185

GAA GCT TCT CGT GCA GCT AAA AAA GAA CTA GAA GCT AAG TAC CAA AAA          857
Glu Ala Ser Arg Ala Ala Lys Lys Glu Leu Glu Ala Lys Tyr Gln Lys
            190                 195                 200

TTA GAG ACT GAT CAC CAA GCC CTA GAA GCT AAG CAC CAA AAA TTA GAG          905
Leu Glu Thr Asp His Gln Ala Leu Glu Ala Lys His Gln Lys Leu Glu
        205                 210                 215

GCT GAT TAC CAA GTT TCA GAG ACT AGC CGT AAG GGT CTA AGT CGT GAC          953
Ala Asp Tyr Gln Val Ser Glu Thr Ser Arg Lys Gly Leu Ser Arg Asp
    220                 225                 230

CTT GAA GCG TCT CGT GAA GCT AAT AAG AAG GTT ACA TCT GAG TTA ACA         1001
Leu Glu Ala Ser Arg Glu Ala Asn Lys Lys Val Thr Ser Glu Leu Thr
235                 240                 245                 250

CAA GCA AAA GCT CAA CTC TCA GCG CTT GAA GAA AGT AAG AAA TTA TCA         1049
Gln Ala Lys Ala Gln Leu Ser Ala Leu Glu Glu Ser Lys Lys Leu Ser
                255                 260                 265

GAA AAA GAA AAA GCT GAG TTA CAA GCA AAA CTA GAT GCA CAA GGA AAA         1097
Glu Lys Glu Lys Ala Glu Leu Gln Ala Lys Leu Asp Ala Gln Gly Lys
            270                 275                 280

GCC CTC AAA GAA CAA TTA GCA AAA CAA ACT GAA GAG CTT GCA AAA CTA         1145
Ala Leu Lys Glu Gln Leu Ala Lys Gln Thr Glu Glu Leu Ala Lys Leu
        285                 290                 295

AGA GCT GAA AAA GCG GCA GGT TCA AAA ACA CCT GCT ACC AAA CCA GCT         1193
Arg Ala Glu Lys Ala Ala Gly Ser Lys Thr Pro Ala Thr Lys Pro Ala
    300                 305                 310

AAT AAA GAA AGA TCA GGT AGA GCT GCT CAA ACA GCT ACA AGA CCT AGC         1241
Asn Lys Glu Arg Ser Gly Arg Ala Ala Gln Thr Ala Thr Arg Pro Ser
```

-continued

```
315                 320                 325                 330

CAA AAT AAA GGA ATG AGA TCA CAA TTA CCG TCA ACA GGC GAA GCA GCT      1289
Gln Asn Lys Gly Met Arg Ser Gln Leu Pro Ser Thr Gly Glu Ala Ala
                335                 340                 345

AAC CCA TTC TTT ACA GCA GCA GCT GCA ACA GTG ATG GTA TCT GCT GGT      1337
Asn Pro Phe Phe Thr Ala Ala Ala Ala Thr Val Met Val Ser Ala Gly
            350                 355                 360

ATG CTT GCT CTA AAA CGC AAA GAA GAA AAC TAAGTCTTTA GAACTTGGTT        1387
Met Leu Ala Leu Lys Arg Lys Glu Glu Asn
        365                 370

TTTGTAACGG TGCAATAGAC AAAAGCAAGC AAGGCCAAAA ACTGAGAAAG TCCTAAAAAG    1447

CTGGCCTTTA CCCCTAAAAA TTAATGTTTT ATAATAAAGA TGTTAGTAAT ATAATTGATA    1507

AATGAGATAC ATTTAATCAT TATGGCAAAA GCAAGAAAAA TAGCTGTATC ATA           1560
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Ser Ala Gly Met Leu Ala Leu Lys Arg Lys Glu Glu Asn
  1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Ala Arg Gln Gln Thr Lys Lys Asn Tyr Ser Leu Arg Lys Leu Lys
  1               5                  10                  15

Thr Gly Thr Ala Ser Val Ala Val Ala Leu Thr Val Leu Gly Ala Gly
            20                  25                  30

Phe Ala Asn Gln Thr Glu Val Arg Ala Asp Glu Ala Lys Lys Met Glu
        35                  40                  45

Val Lys Glu Ser Glu Lys Glu Ser Gln Tyr Lys Thr Leu Ala Leu Arg
    50                  55                  60

Gly Glu Asn Ala Asp Leu Arg Asn Val Asn Ala Lys Tyr Leu Glu Lys
65                  70                  75                  80

Ile Asn Ala Glu Glu Glu Lys Asn Lys Lys Leu Glu Ala Ile Asn Lys
                85                  90                  95

Glu Leu Asn Glu Asn Tyr Tyr Lys Leu Gln Asp Gly Ile Asp Ala Leu
            100                 105                 110

Glu Lys Glu Lys Glu Asp Leu Lys Thr Thr Leu Ala Lys Thr Thr Lys
        115                 120                 125

Glu Asn Glu Ile Ser Glu Ala Ser Arg Lys Gly Leu Ser Arg Asp Leu
    130                 135                 140

Glu Ala Ser Arg Thr Ala Lys Lys Glu Leu Glu Ala Lys His Gln Lys
145                 150                 155                 160

Leu Glu Ala Glu Asn Lys Lys Leu Thr Glu Gly Asn Gln Val Ser Glu
                165                 170                 175

Ala Ser Arg Lys Gly Leu Ser Asn Asp Leu Glu Ala Ser Arg Ala Ala
            180                 185                 190
```

-continued

Lys	Lys	Glu	Leu	Glu	Ala	Lys	Tyr	Gln	Lys	Leu	Glu	Thr	Asp	His	Gln
		195					200					205			
Ala	Leu	Glu	Ala	Lys	His	Gln	Lys	Leu	Glu	Ala	Asp	Tyr	Gln	Val	Ser
	210					215					220				
Glu	Thr	Ser	Arg	Lys	Gly	Leu	Ser	Arg	Asp	Leu	Glu	Ala	Ser	Arg	Glu
225					230					235					240
Ala	Asn	Lys	Lys	Val	Thr	Ser	Glu	Leu	Thr	Gln	Ala	Lys	Ala	Gln	Leu
				245					250					255	
Ser	Ala	Leu	Glu	Glu	Ser	Lys	Lys	Leu	Ser	Glu	Lys	Glu	Lys	Ala	Glu
			260					265					270		
Leu	Gln	Ala	Lys	Leu	Asp	Ala	Gln	Gly	Lys	Ala	Leu	Lys	Glu	Gln	Leu
		275					280					285			
Ala	Lys	Gln	Thr	Glu	Glu	Leu	Ala	Lys	Leu	Arg	Ala	Glu	Lys	Ala	Ala
	290					295					300				
Gly	Ser	Lys	Thr	Pro	Ala	Thr	Lys	Pro	Ala	Asn	Lys	Glu	Arg	Ser	Gly
305					310					315					320
Arg	Ala	Ala	Gln	Thr	Ala	Thr	Arg	Pro	Ser	Gln	Asn	Lys	Gly	Met	Arg
				325					330					335	
Ser	Gln	Leu	Pro	Ser	Thr	Gly	Glu	Ala	Ala	Asn	Pro	Phe	Phe	Thr	Ala
			340					345					350		
Ala	Ala	Ala	Thr	Val	Met	Val	Ser	Ala	Gly	Met	Leu	Ala	Leu	Lys	Arg
		355					360					365			
Lys	Glu	Glu	Asn												
	370														

We claim:

1. An isolated DNA molecule of DNA encoding the protein represented by SEQ ID NO. 3.

2. A DNA molecule, according to claim **1**, wherein said molecule is represented by SEQ ID NO. 1.

3. A recombinantly produced plasmid comprising DNA represented by SEQ ID NO. 1.

4. A plasmid according to claim **3**, wherein said plasmid is pML2-14.

5. A cell in culture transformed with the DNA of claim **1**.

6. A cell according to claim **5**, wherein said cell is *Escherichia coli.*

7. A process for preparing an IgA-binding polypeptide comprising

cultivating the transformed cell of claim **5** under conditions such that said polypeptide is expressed, and

recovering said polypeptide from the culture.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 5,352,588

DATED : October 4, 1994

INVENTOR(S) : Fischetti et al.

It is certified that error appears in the above-indentified patent and that said Letters Patent is hereby corrected as shown below:

Column 2, line 28: Delete "FIG. 1" and insert --SEQ ID NO. 1--

Column 3, line 7: Delete "the an may be used." and insert --the art may be used.--

Signed and Sealed this

Seventh Day of February, 1995

Attest:

BRUCE LEHMAN

Attesting Officer

Commissioner of Patents and Trademarks